Making Space for Governance

Sarah Clarke - Owner Infospectives Ltd

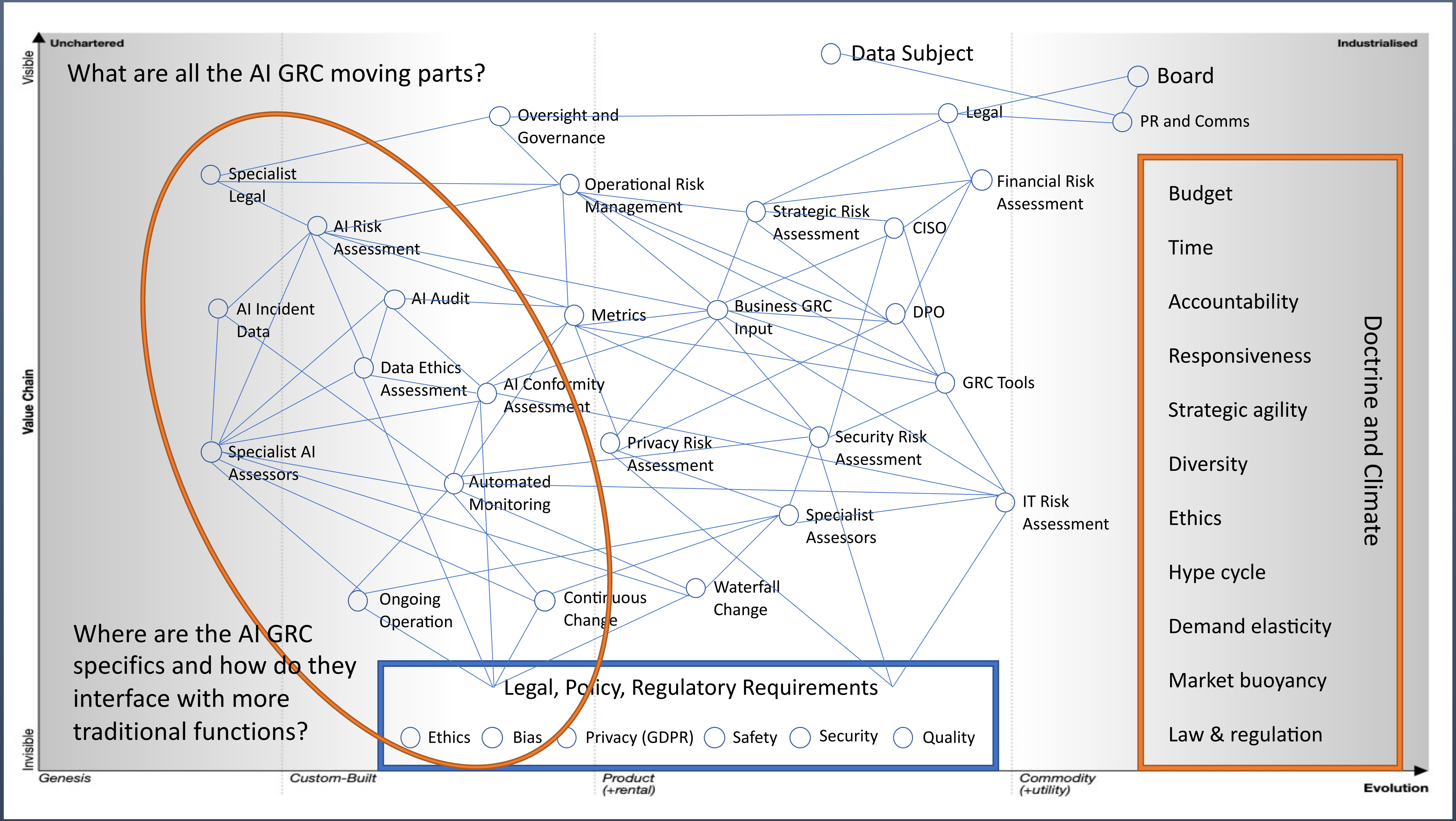
Cybersecurity, Data Protection, Novel Technology GRC

Vendor Security Governance Guest Lecturer for Manchester University, IEEE and World Ethical Data Foundation Contributor, Emeritus Fellow ForHumanity

Infospectives™

“No-one should be accountable for something they can’t influence, or don’t understand”

Lines of Communication



Infospectives™

In a nutshell:

Accountability for risks must be FORMALLY assigned to decision-makers who have influence and means to effect change…

…decision-makers must REMAIN accountable until data/tech use for specified purposes ceases, or the role is formally handed over…

…SPECIALISTS are accountable for providing clear information about requirements, risks, and blockages…

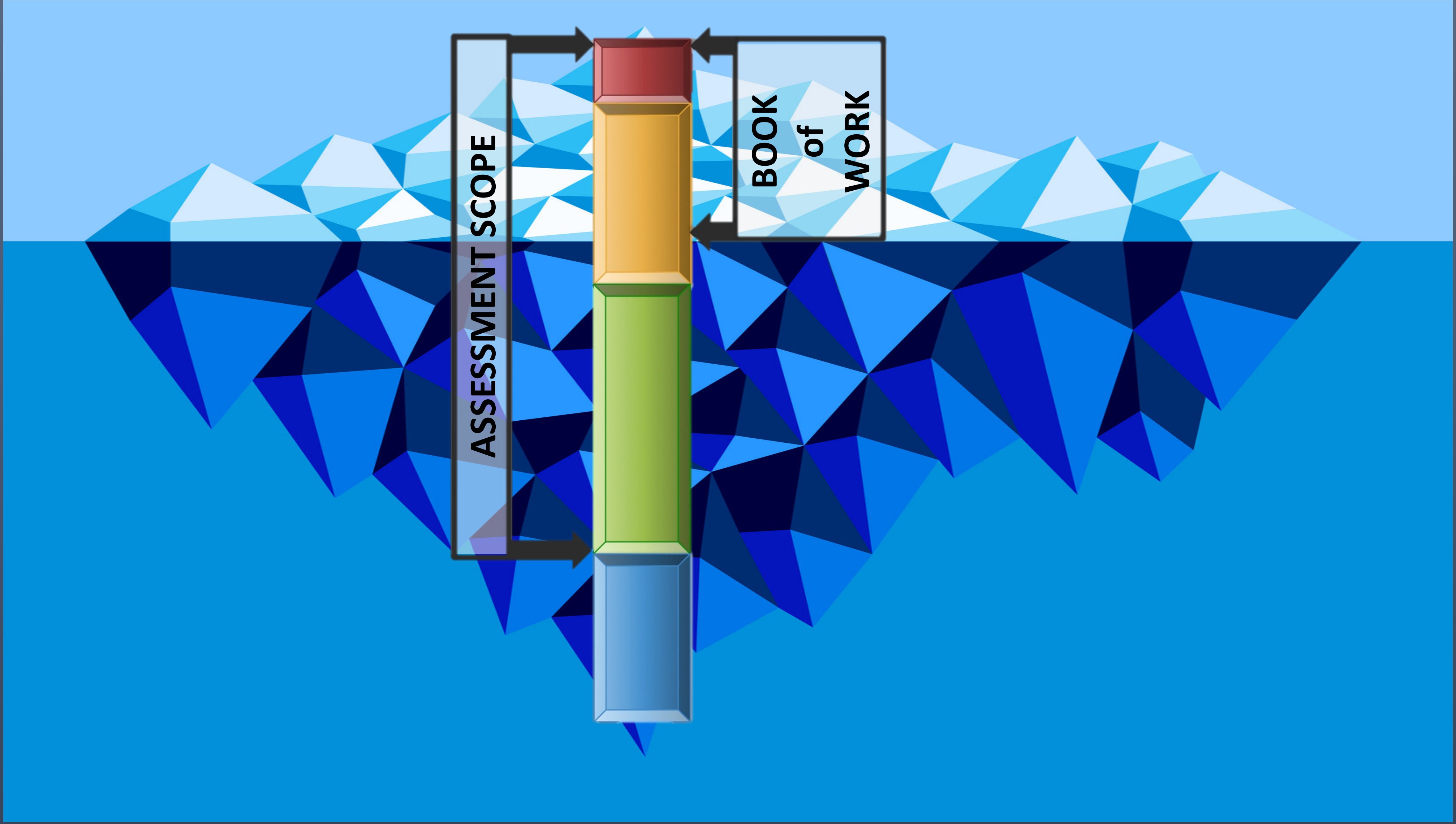
DECISION-MAKERS are accountable for providing sufficient time, money, and support to make that work...

…because NO-ONE should be accountable for something that they can’t influence, or don’t understand.

“Being uncertain is not the same as being at risk…

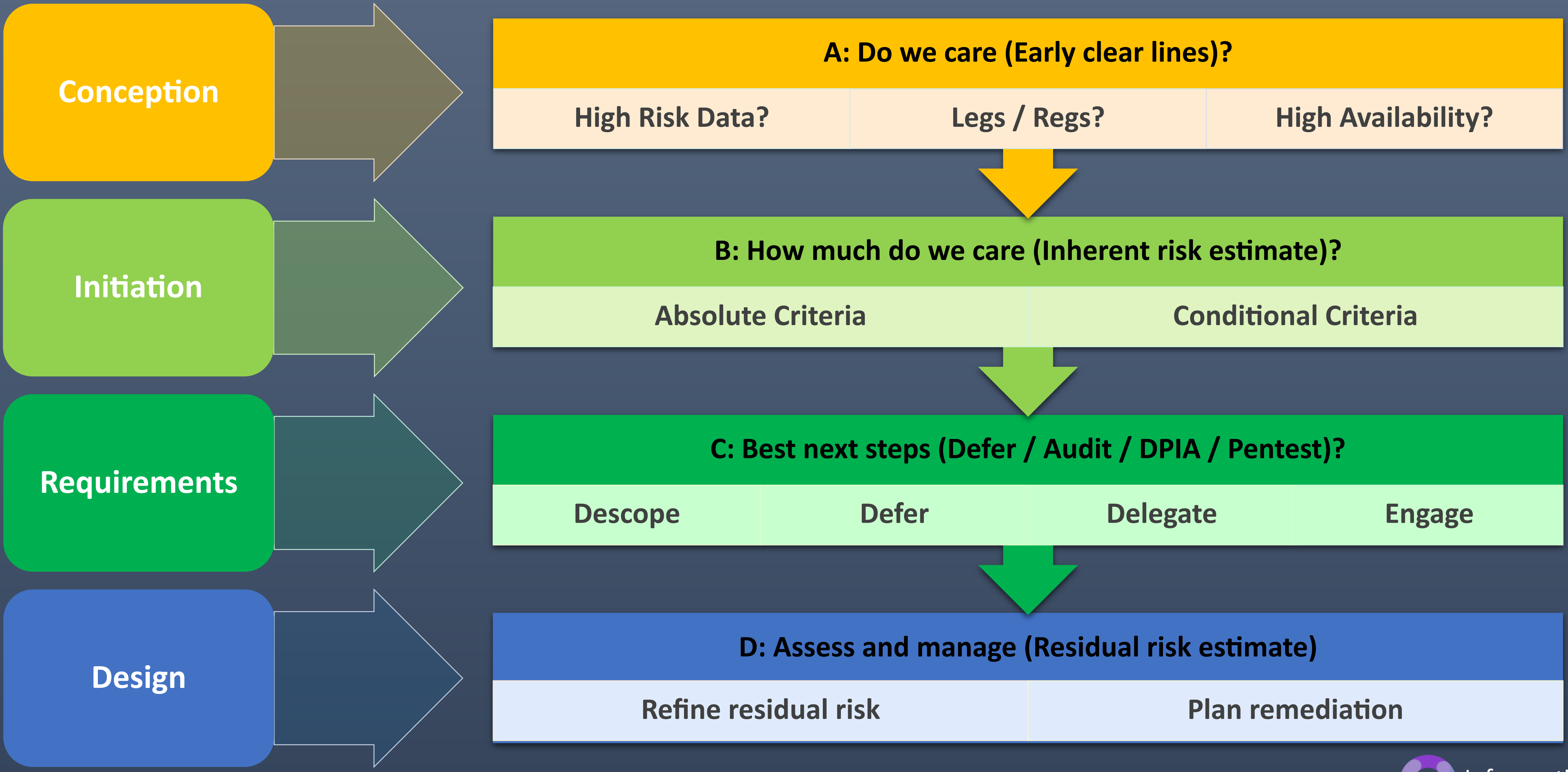
…and being at risk is not the same as being at intolerable or proximate risk”

Early Triage



Infospectives™

Where would you start?



Conception

Initiation

Requirements

Design

A: Do we care (Early clear lines)?

High Risk Data?	Legs / Regs?	High Availability?

B: How much do we care (Inherent risk estimate)?

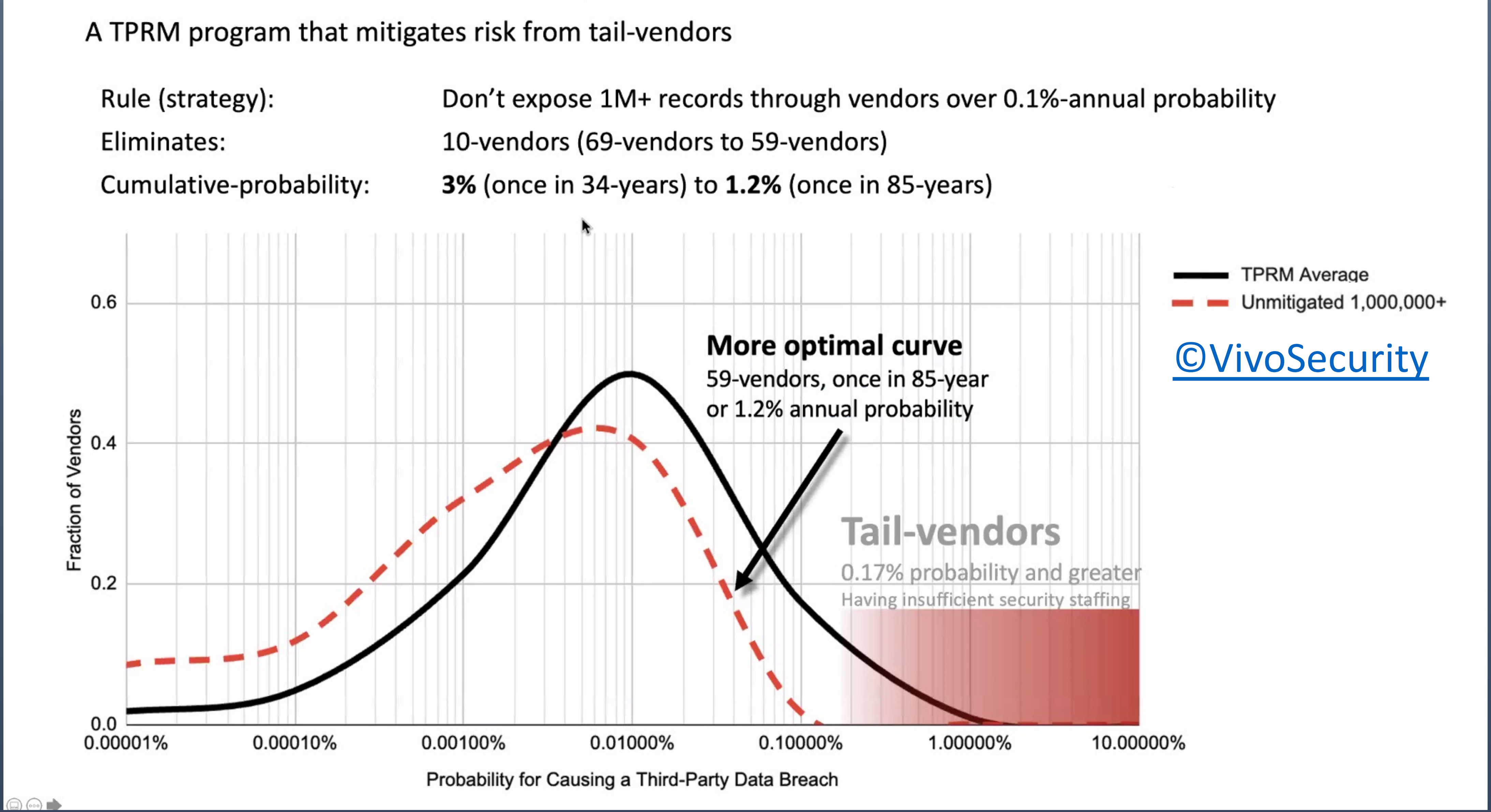
Absolute Criteria	Conditional Criteria

C: Best next steps (Defer / Audit / DPIA / Pentest)?

Descope	Defer	Delegate	Engage

D: Assess and manage (Residual risk estimate)

Refine residual risk	Plan remediation



A TPRM program that mitigates risk from tail-vendors

Rule (strategy):	Don’t expose 1M+ records through vendors over 0.1%-annual probability
Eliminates:	10-vendors (69-vendors to 59-vendors)
Cumulative-probability:	**3%** (once in 34-years) to **1.2%** (once in 85-years)

More optimal curve
59-vendors, once in 85-year
or 1.2% annual probability

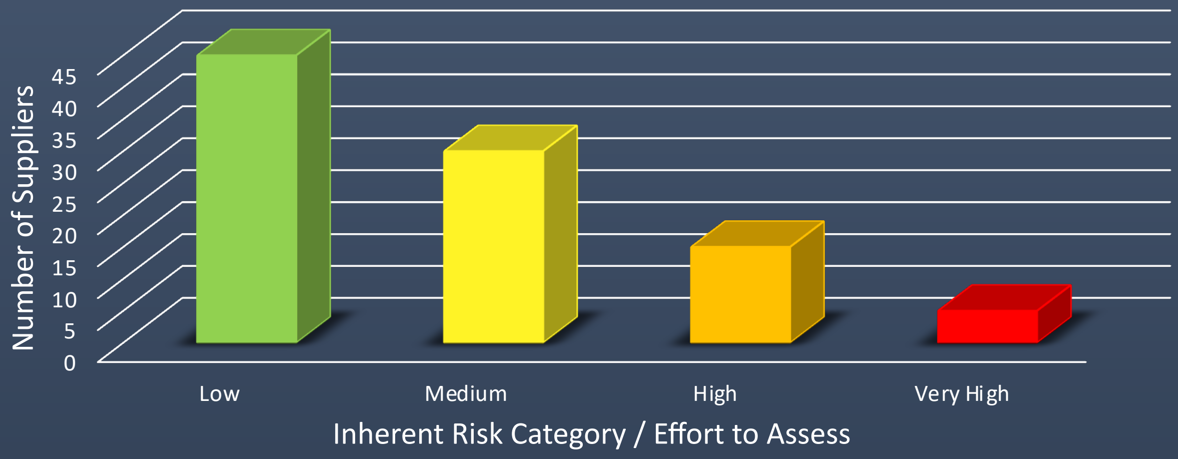
Tail-vendors
0.17% probability and greater
Having insufficient security staffing

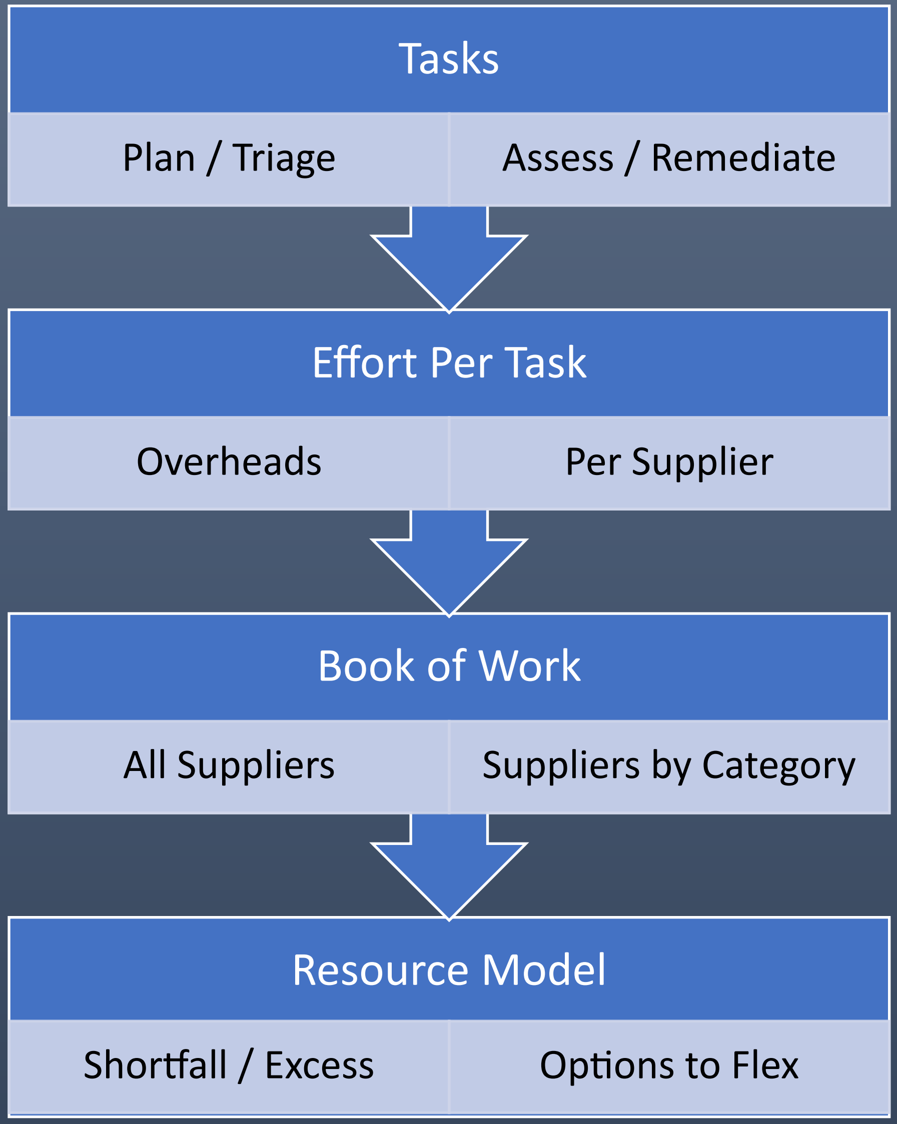
©VivoSecurity

“The earlier and more constructively you escalate, the less it sounds like an excuse”

Resource modelling

Activity	Per Supplier / Overhead
Triage	Per Supplier
Assessment	Per Supplier
Remediation / Retesting	Per Supplier
Contract negotiation	Per Supplier
Governance data collation	Per Supplier
Remediation data collation	Per Supplier
Regular on-going governance	Per Supplier
Risk management	Per Supplier
Metrics and reporting	Overhead
Process development and planning	Overhead
Training	Overhead
Stakeholder management	Overhead





Tasks

Plan / Triage | Assess / Remediate

Effort Per Task

Overheads | Per Supplier

Book of Work

All Suppliers | Suppliers by Category

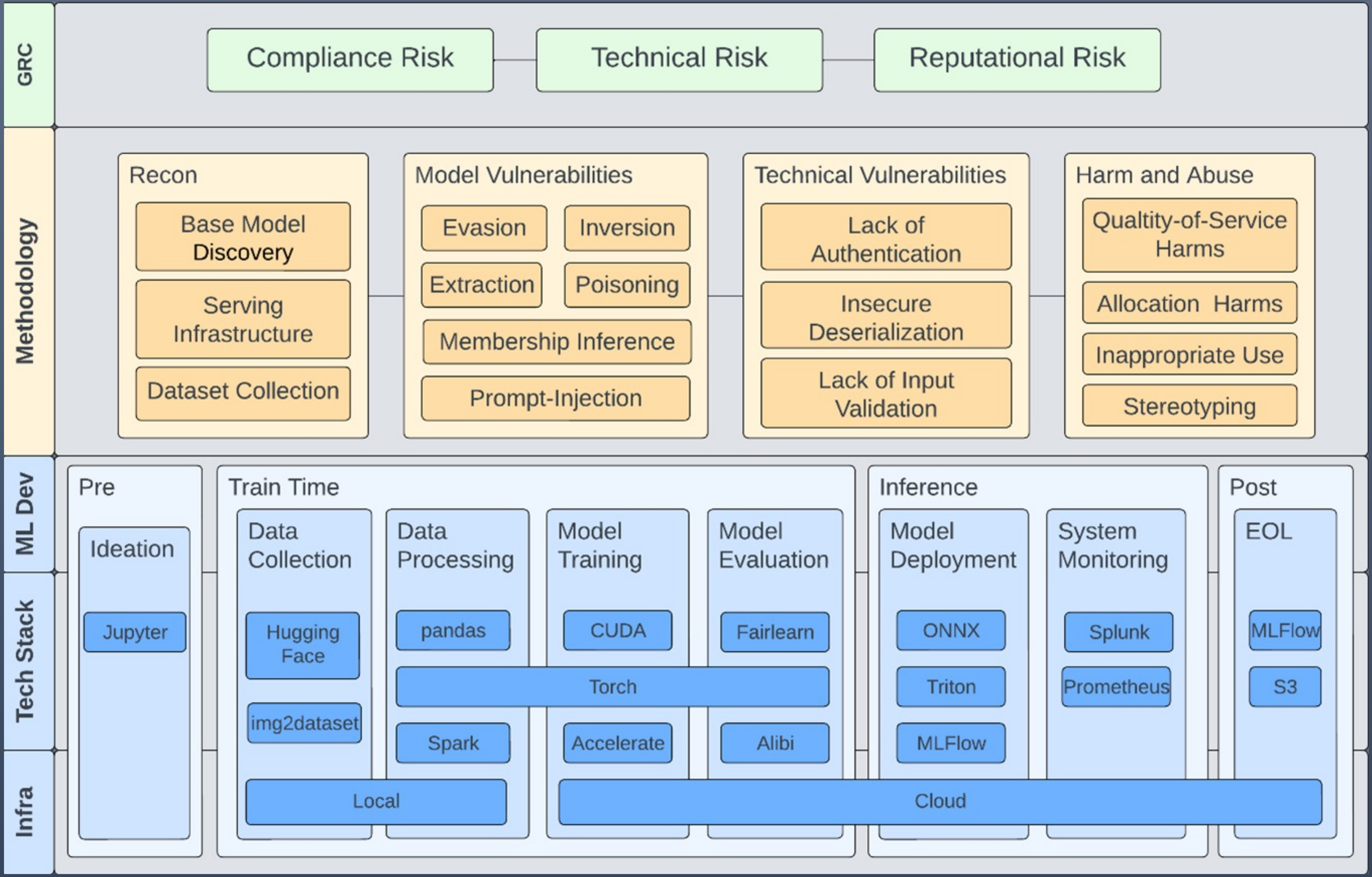
Resource Model

Shortfall / Excess | Options to Flex

Infospectives™

“Uncertainty is not a showstopper, unless you have no plan to cover bets”

Defining uncertainty



GRC

Compliance Risk — Technical Risk — Reputational Risk

Methodology

- Recon
  - Base Model Discovery
  - Serving Infrastructure
  - Dataset Collection
- Model Vulnerabilities
  - Evasion
  - Inversion
  - Extraction
  - Poisoning
  - Membership Inference
  - Prompt-Injection
- Technical Vulnerabilities
  - Lack of Authentication
  - Insecure Deserialization
  - Lack of Input Validation
- Harm and Abuse
  - Qualtity-of-Service Harms
  - Allocation Harms
  - Inappropriate Use
  - Stereotyping

ML Dev

- Pre
  - Ideation
- Train Time
  - Data Collection
  - Data Processing
  - Model Training
  - Model Evaluation
- Inference
  - Model Deployment
  - System Monitoring
- Post
  - EOL

Tech Stack

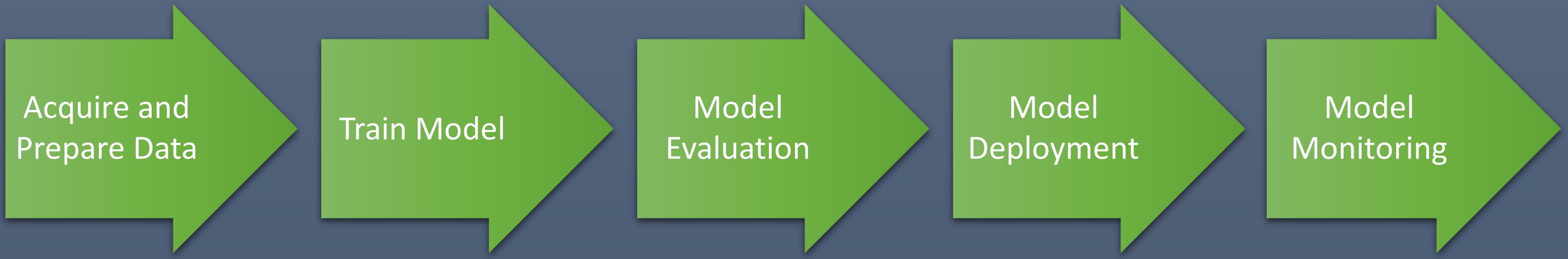
- Ideation: Jupyter
- Data Collection: Hugging Face, img2dataset
- Data Processing: pandas, Spark
- Model Training: CUDA, Accelerate
- Model Evaluation: Fairlearn, Alibi
- Data Processing to Model Evaluation: Torch
- Model Deployment: ONNX, Triton, MLFlow
- System Monitoring: Splunk, Prometheus
- EOL: MLFlow, S3

Infra

- Local
- Cloud

Source: NVIDIA Red Teaming Introduction, June 2023

AI LIFECYCLE



AI AUDIT / ASSESSMENT – OVERSIGHT / DOCUMENTATION / PROCESS - REVIEW

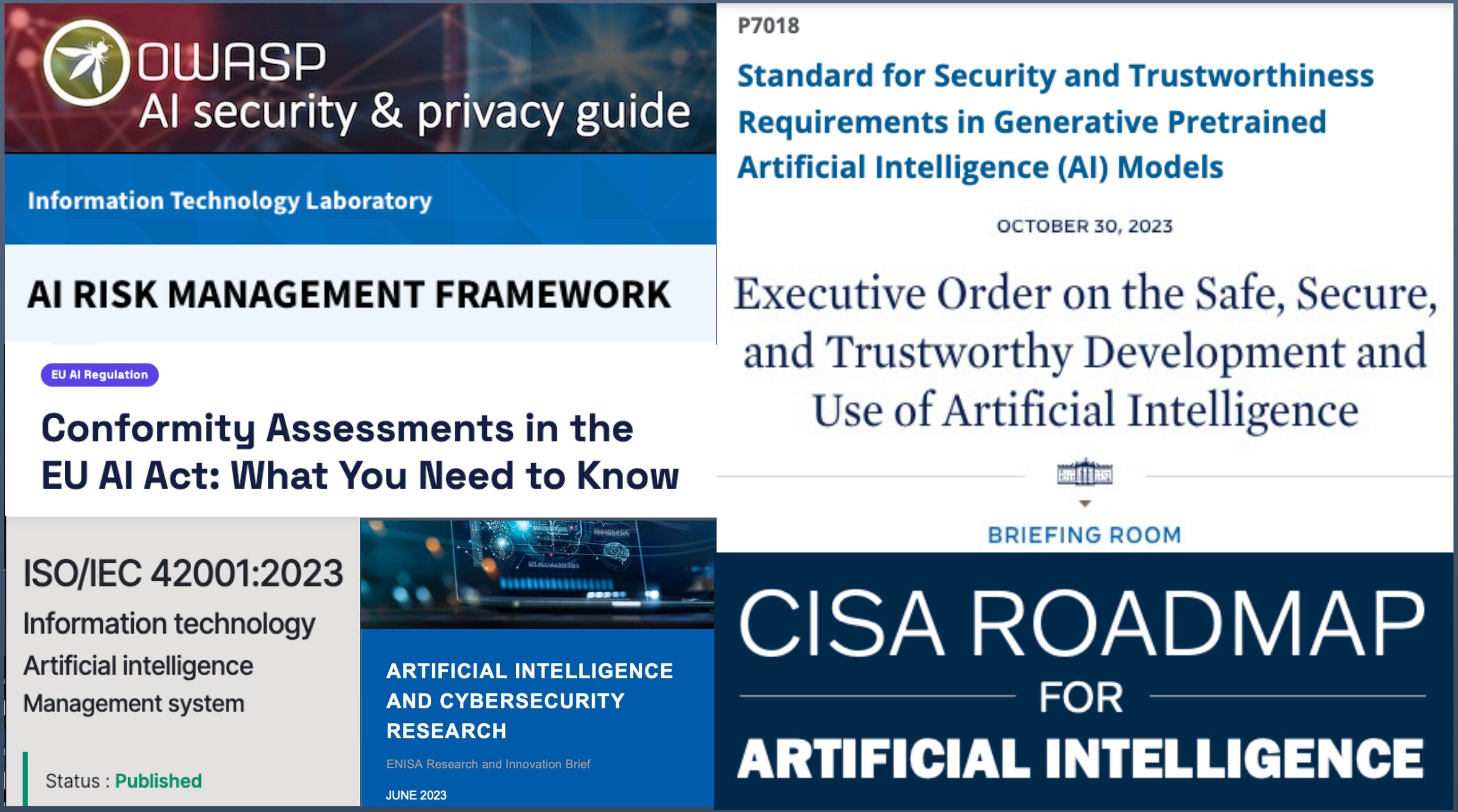
Compliant	Compliant	Compliant	Compliant	Compliant
Non-Compliant	Non-Compliant	Non-Compliant	Non-Compliant	Non-Compliant
Unknown	Unknown	Unknown	Unknown	Unknown

AI AUDIT / ASSESSMENT – SYSTEM / CODE / DATA - TECHNICAL TESTING

Compliant	Compliant	Compliant	Compliant	Compliant
Non-Compliant	Non-Compliant	Non-Compliant	Non-Compliant	Non-Compliant
Unknown	Unknown	Unknown	Unknown	Unknown

Infospectives ™

Assurance and governance guidance is evolving



OWASP
AI security & privacy guide

Information Technology Laboratory

AI RISK MANAGEMENT FRAMEWORK

EU AI Regulation

Conformity Assessments in the EU AI Act: What You Need to Know

ISO/IEC 42001:2023
Information technology
Artificial intelligence
Management system

Status : Published

ARTIFICIAL INTELLIGENCE AND CYBERSECURITY RESEARCH
ENISA Research and Innovation Brief
JUNE 2023

CISA ROADMAP FOR ARTIFICIAL INTELLIGENCE

BRIEFING ROOM

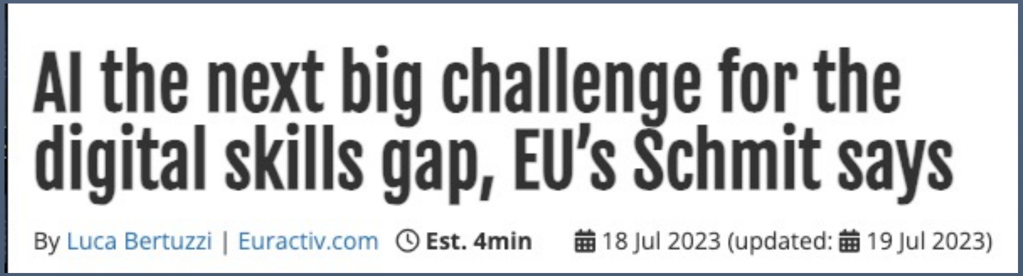
Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

OCTOBER 30, 2023

P7018
Standard for Security and Trustworthiness Requirements in Generative Pretrained Artificial Intelligence (AI) Models

Sources counter-clockwise: OWASP, NIST, HolisticAI, ISO/IEC, ENISA, CISA, The White House, IEEE, “You can’t recruit for deep local knowledge”

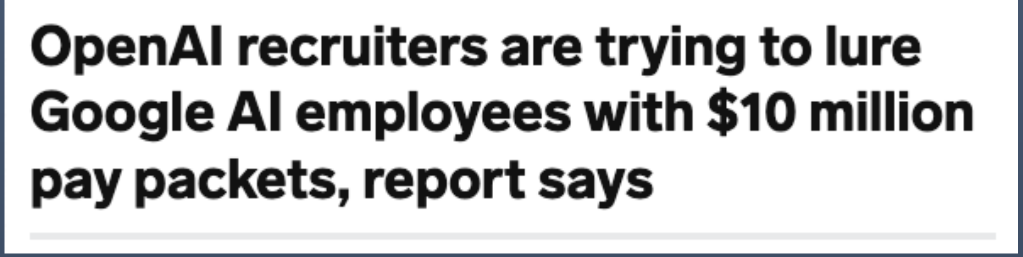
Live the hiring rhetoric



AI the next big challenge for the digital skills gap, EU's Schmit says

By Luca Bertuzzi | Euractiv.com Est. 4min 18 Jul 2023 (updated: 19 Jul 2023)

Source: Euroactive 2023



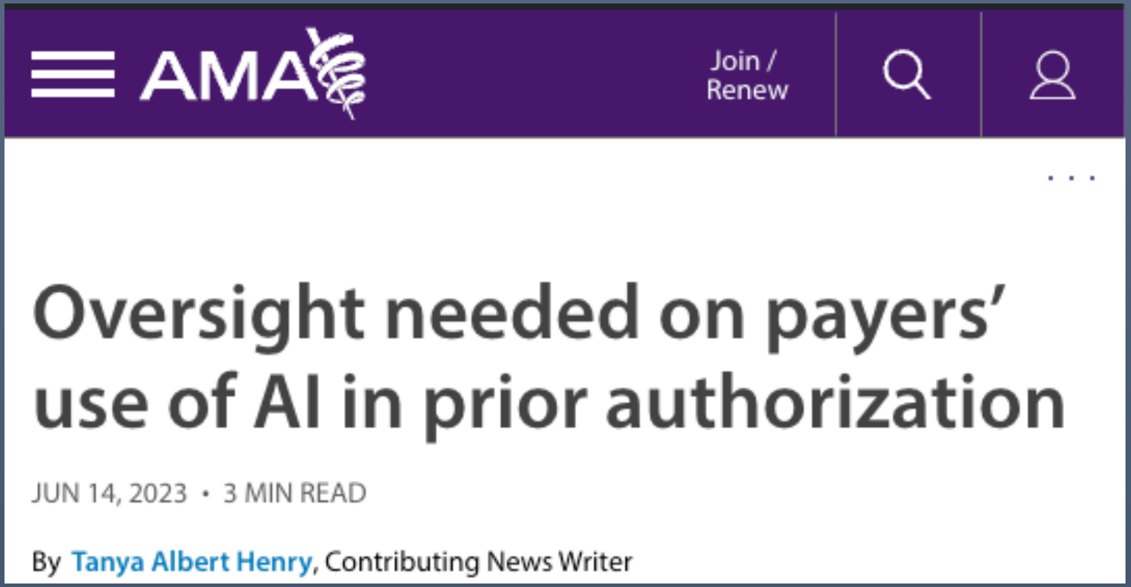
OpenAI recruiters are trying to lure Google AI employees with $10 million pay packets, report says

Source: Business Insider 2023

“It is about adverse outcomes from accidents, incidents, and breaches…

…plus potential impact when systems work exactly as we said they would”

Health Insurance AI - What is the risk and downstream Impact?



Oversight needed on payers' use of AI in prior authorization

JUN 14, 2023 • 3 MIN READ

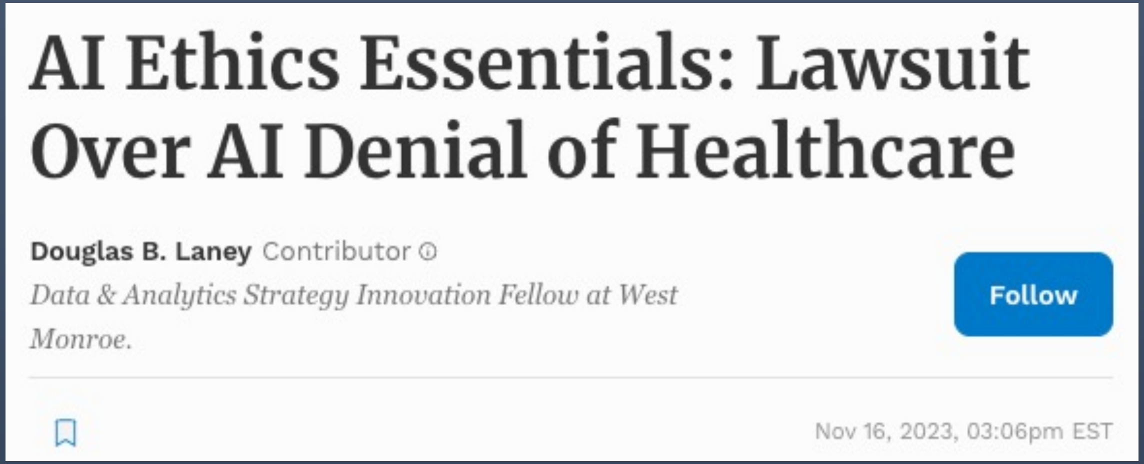
By Tanya Albert Henry, Contributing News Writer

Source: American Medical Association, June 2023

"…doctors spending an average of 1.2 seconds on each case,"

"enabling… 'fast, efficient and streamlined coverage decisions.'"

"…there's a fine line between effective oversight and misconceptions of the technology… Even well-intentioned governance can be hampered by a lack of technical knowledge" Ryan Elmore



AI Ethics Essentials: Lawsuit Over AI Denial of Healthcare

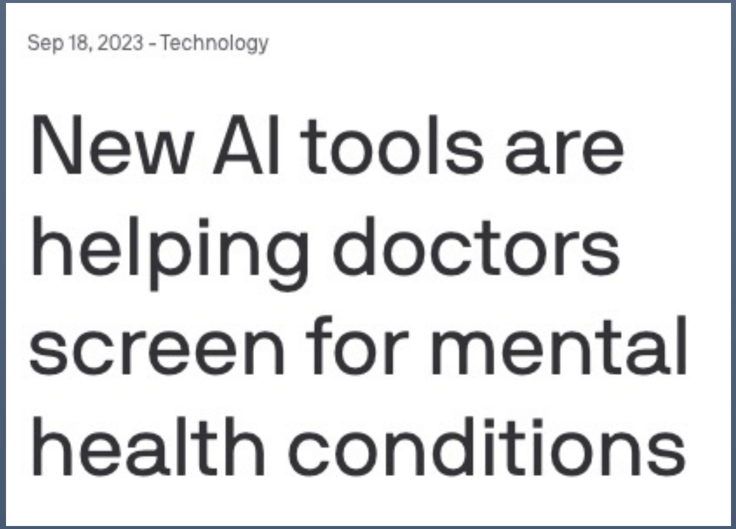
Douglas B. Laney Contributor ⓘ

Data & Analytics Strategy Innovation Fellow at West Monroe.

Nov 16, 2023, 03:06pm EST

Source: Forbes, November 2023

Are health triage chatbots adequately governed?



Sep 18, 2023 - Technology

New AI tools are helping doctors screen for mental health conditions

Source: Axios, September 2023

"There are not enough trained mental health professionals on the planet to serve the astronomical disease prevalence,"
Ross Harper, Limbic AI

“…data engineering for AI models seems to be overlooked or misunderstood… shortcomings may indicate overly accelerated promotion”
Dr Novillo-Ortiz, WHO Europe Unit Head, Data and Digital Health

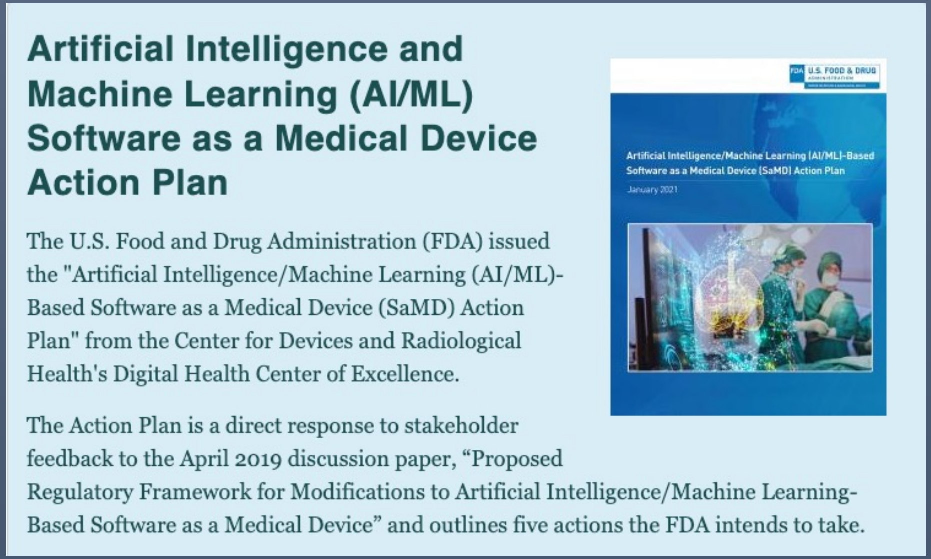
©

Artificial intelligence in mental health research: new WHO study on applications and challenges

6 February 2023 | News release | Reading time: 3 min (678 words)

Source: World Health Organisation, February 2023

Progress governing Software as a Medical Device (SaMD)



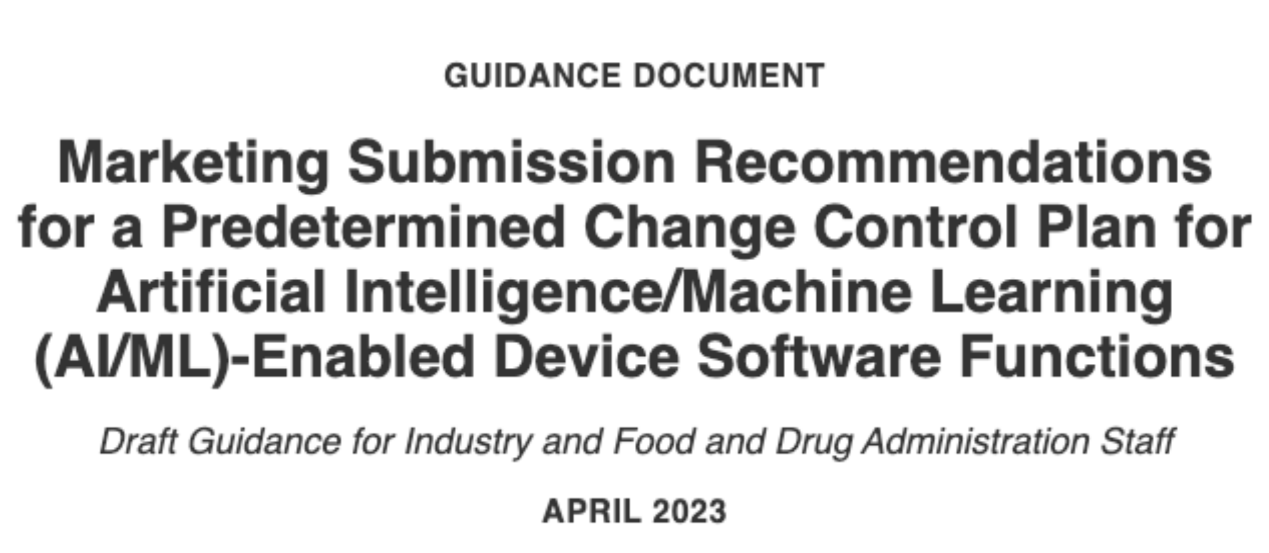
Artificial Intelligence and Machine Learning (AI/ML) Software as a Medical Device Action Plan

The U.S. Food and Drug Administration (FDA) issued the "Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan" from the Center for Devices and Radiological Health's Digital Health Center of Excellence.

The Action Plan is a direct response to stakeholder feedback to the April 2019 discussion paper, "Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning-Based Software as a Medical Device" and outlines five actions the FDA intends to take.

U.S. FOOD & DRUG ADMINISTRATION
Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan
January 2021

Source: FDA, January 2021



GUIDANCE DOCUMENT

Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence/Machine Learning (AI/ML)-Enabled Device Software Functions

Draft Guidance for Industry and Food and Drug Administration Staff

APRIL 2023

Source: FDA, April 2023

DECEMBER 14, 2023

Delivering on the Promise of AI to Improve Health Outcomes

Source: The White House, December 2023

"The commitments received today will serve to align industry action on AI around the "FAVES" principles—that AI should lead to healthcare outcomes that are Fair, Appropriate, Valid, Effective, and Safe"

Source: The White House, December 2023

Getting In touch

Sarah.Clarke@Infospectives.co.uk

Linkedin.com/in/infospectives

Noted resources, concepts and organizations

Third-party items on slides have links to sources. Other resources in rough order mentioned:

- The World Ethical Data Foundation
- For Humanity
- Six Sigma Gemba and Gemba walks
- Cynefin Entangled Trios and other complexity management resources
- Wardley Mapping visualization method for complex systems
- Sherri Douville ‘Advanced Health Technology’ and other publications
- EU AI Act prohibited and high-risk AI classification
- OWASP OpenCRE and Top 10 for LLMs (Large Language Models)
- IEEE P3119 Draft AI Procurement Standard
- Vivo Security statistical analysis of data breaches and application to Third Party Risk Management
- Factor Analysis of Information Risk (FAIR), cybersecurity risk modeling and quantification
- Retrieval Augmented Generation (RAG) for LLMs
- Executive Order to Protect Americans’ Sensitive Personal Data